Object Detection in Images Based on Homogeneous Region Segmentation

Abdesalam Amrane[1,2(✉)], Abdelkrim Meziane[1],
and Nour El Houda Boulkrinat[1]

[1] Research Center on Scientific and Technical Information (CERIST),
16030 Ben Aknoun, Algiers, Algeria
`amrane@mail.cerist.dz`
[2] Université Abderrahmane Mira Béjaia,
Rue Targa Ouzemour, 06000 Béjaia, Algeria

Abstract. Image segmentation for object detection is one of the most fundamental problems in computer vision, especially in object-region extraction task. Most popular approaches in the segmentation/object detection tasks use sliding-window or super-pixel labeling methods. The first method suffers from the number of window proposals, whereas the second suffers from the over-segmentation problem. To overcome these limitations, we present two strategies: the first one is a fast algorithm based on the region growing method for segmenting images into homogeneous regions. In the second one, we present a new technique for similar region merging, based on a three similarity measures, and computed using the region adjacency matrix. All of these methods are evaluated and compared to other state-of-the-art approaches that were applied on the Berkeley image database. The experimentations yielded promising results and would be used for future directions in our work.

Keywords: Region proposal · Region growing · Region merging
Image segmentation

1 Introduction

Object detection is the process of finding instances of semantic objects such as persons, animals and vehicle in images or videos. One way to address the object detection problem is to use image segmentation methods. Currently, many applications require an automatic segmentation step in very different fields, with specificities related to the processed images [1]. Image segmentation is an essential process for many applications, such as object recognition, target tracking, content-based image retrieval [2]. Image segmentation methods are mainly grouped into three categories: Threshold-based techniques, Edge-based techniques and Region-based techniques [3].

More recently, hybrid models, which incorporate both edge and region information in their segmentation algorithms, have been introduced [4]. Image segmentation can be classified into supervised and unsupervised segmentation methods [5]. In this work, we are interested in unsupervised methods (automated segmentation methods).

The unsupervised segmentation algorithms divides the image into homogeneous regions (a set of neighboring pixels) but without considering the significance of these regions. While human segmentation tends to divide the image into objects that have a meaning (called semantic segmentation). For this reason, researchers try to bring automatic segmentation as close as possible to human segmentation.

In this paper, we propose a fully automated approach to region segmentation in images, that combines region growing and edge preservation methods. We also give a brief description of our implementation and evaluation results obtained using the Berkeley Segmentation Dataset[1] (BSD).

2   Related Works

Object detection can be treated in two phases. The first phase uses the region proposals method to segment an image into candidate objects. The second phase classifies every proposal region into different classes of objects using trained classifiers such as SVM [6] or Neural Network [7]. In object detection, some of the efficient techniques exploit sliding window and boosting [8]. The main drawback of the sliding window is that the number of proposals can have a complexity of $O(10^6)$ for a 640×480 image [9], which increases the computational time and cost. To solve these problems, the authors of [9] proposed to use the superpixel labeling method to effectively detect objects. Superpixel represents a set of pixels which have similar colors and spatial relationships. Unfortunately, the use of superpixels also suffers from the problem of choosing the number of clusters K (maximal number of objects present in an image). A small value will generate a sub-segmentation and a great value will generate an over-segmentation.

There are several image segmentation techniques that were proposed in the literature. They are classified in three categories: Threshold Technique Segmentation, Region Based Segmentation and Edge Based Segmentation. Each one of these approaches has advantages and disadvantages depending on the application domains [10,11]. Recently, the authors of [1] proposed an interested work based on region merging strategies that start by an over-segmentation process. Unfortunately, the author does not quote the execution time of the segmentation process.

We present in this work, an unsupervised image segmentation method for object detection. The proposed approach uses the region growing method combined with edge detection and some filters like bilateral filter, which serves to smooth images.

---

[1] http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html.

3   Proposed Method

In this work, we propose a new hybrid segmentation method that combines two techniques: edge based region growing and region merging method for object detection. For this reason, we use our algorithm which is based on region growing method (pixel aggregation) reinforced using canny edge to preserve the boundaries of objects. The generic algorithm is composed of three steps. we present the steps of our method Fig. 1 which are detailed in the following subsections.

Fig. 1. Architecture of proposed method

3.1   Image Pre-processing

– Denoise image using median filter,
– Compute canny edge for original image I,
– Smooth original image I (using bilateral filter),
– Convert smoothed image to LAB color space,
– Normalize color values between 0 and 1.

3.2   Region Growing Segmentation

– Select initial seed point located in position (1,1) from image I,
– Define a similarity measure S(p,q) based on two criteria:
   • Euclidean distance between the LAB color of pixels p and q,
   • Verify non-edge pixels for pixel p and q.
– Evaluate the neighbors of seed point (seed$_p$) as above:
   • If the neighboring pixels of seed point satisfies the defined criteria, they will be grown (add pixel q to seed$_p$). The 3 neighbors of pixel q are added into the neighbors list of p according to the moving direction between pixels p and q,
   • Else consider pixel q as a new seed point (seed$_q$).
– Repeat the process until all pixels in the image are treated.

We use 8-connected neighborhood to grow the neighboring pixels to the current seed. The formula of Euclidian distance used is:

$$D_{Lab}(x, y) = \sqrt{(L_x - L_y)^2 + (a_x - a_y)^2 + (b_x - b_y)^2} \tag{1}$$

We grow each pixel that has distance value D$_{Lab}$ less than or equal to the best threshold; I choose value 0.10 (that is perceptually acceptable). To examine the neighboring pixels, we enumerate each neighboring pixels of a given pixel p in a clockwise direction {0,1,..,7} which corresponds to the orientation degree of pixels. This technique improves the processing time and allows a fast scan of the whole image.

3.3   Region Merging

For an initial-segmented image:

– Define a criterion for merging two adjacent regions,
– Compute features for each region and construct a region adjacency matrix,
– Iteratively merge all adjacent regions satisfying the merging criterion,
– Repeat process until all regions are merged.

The merging criterions that we propose are the following:

– Regions are allowed to merge if they have small color differences (using pre-
  vious defined Euclidian distance and mean colors of each region),
– Region that have small size will be merged with the biggest adjacent region,
– All adjacent regions with Lambda value less than a given threshold will be
  merged.

We use Full Lambda Schedule method introduced by Robinson [12], described
as follow:

$$L\left(v_i, v_j\right) = \frac{\frac{|O_i|.|O_j|}{|O_i|+|O_j|}.||u_i - u_j||^2}{l(v_i, v_j)} \tag{2}$$

where:

$|O_i|$ and $|O_j|$ are the areas of regions i and j,
$||u_i\text{-}u_j||^2$ is the euclidean distance between the mean color values of regions i and j,
$l(v_j, v_t)$ is the length of the common boundary of regions i and j.

4   Experimental Evaluation

In order to make an objective comparison between different segmentation meth-
ods, we use some evaluation criteria wich have already been defined in literature.
Briefly stated, there are two main approaches [13]: supervised evaluation criteria
and unsupervised evaluation criteria. Supervised evaluation use a ground truth,
whereas unsupervised evaluation enable the quantification quality of a segmen-
tation result without any prior knowledge [14]. The evaluation of a segmentation
result makes sense at a given level of precision.

In our work, we choose to use a supervised evaluation. Our evaluation is
based on the Berkeley Segmentation Dataset and for each of these algorithms,
we examine two metrics:

1. Probabilistic Rand Index: measures the probability that the pair of samples
   have consistent labels in the two segmentations. The range of PRI is between
   [0,1], with larger value indicating greater similarity between two segmenta-
   tions [15];
2. Variation of Information: measures how much we can know of one segmen-
   tation given another segmentation [15]. The range of VoI is between $[0, \infty]$,
   smaller value indicating better results.

Table 1. Quantitative comparison of different algorithms

Algorithms	PRI	VoI
Efficient graph-based image segmentation	0.770	2.188
Texture and boundary encoding-based segmentation	0.785	2.002
Weighted modularity segmentation	0.752	2.103
Mean-shift	0.772	2.004
Marker controlled watershed	0.753	2.203
Multiscale normalized cut	0.742	2.651
Our's	0.764	2.191

Table 2. Best results of segmentation process

Original images	Segmentation results	Measures
		PRI = 0.91 VoI = 1.12 Regions = 10
		PRI = 0.93 VoI = 1.68 Regions = 21
		PRI = 0.97 VoI = 0.66 Regions = 7
		PRI = 0.95 VoI = 0.85 Regions = 24

Our experiments were performed on a machine with an intel core i7-2630QM. We compare our method with other segmentation algorithms, including: Efficient Graph-Based Image Segmentation [16], Texture and Boundary Encoding-based Segmentation [17], Weighted Modularity Segmentation [18], Mean-Shift

[19], Marker Controlled Watershed [20], Multiscale Normalized Cut [21]. The segmentation results are presented in Table 1.

Our method for unsupervised image segmentation gives acceptable results especially we have used only color feature for similarity measure. We observe that the average time execution is less than 25 s. As a result, we can say that our method gives a high quality segmentation result with minimum time consumption. Table 2 shows some examples:

5   Conclusions

In this paper, we have proposed a hybrid method for unsupervised image segmentation. We have taken advantage of the fast edge based region growing algorithm and region merging techniques to overcome the problems of over-segmentation and sub-segmentation encountered in the super-pixel and sliding window for object detection.

Our algorithms have been tested on the publicly available Berkeley Segmentation Dataset as well as on the Semantic Segmentation Dataset. Then, they have been compared with other popular algorithms. Experimental results have demonstrated that our algorithms have produced acceptable results. However, the obtained regions do not exactly match the shapes of the objects contained in the images. Other treatment would be required to make the correspondence between regions and objects. Thereby, we can confirm the complexity of the segmentation domain and then we will explore other approaches to improve the object detection process by using machine learning in our future works.

References

1. Krahenbuhl, A.: Segmentation et analyse géométrique: application aux images tomodensitométriques de bois. Thesis, Université de Lorraine (2014)
2. Peng, B., Zhang, L., Zhang, D.: Automatic image segmentation by dynamic region merging. IEEE Trans. Image Process. **20**(12), 3592–3605 (2011)
3. Hedberg, H.: A survey of various image segmentation techniques. Department of Electroscience, Box 118 (2010)
4. Fox, V., Milanova, M., Al-Ali, S.: A hybrid morphological active contour for natural images. Int. J. Comput. Sci. Eng. Appl. **3**(4), 1–13 (2013)
5. Niyas, S., Reshma, P., Thampi, S.M.: A color image segmentation scheme for extracting foreground from images with unconstrained lighting conditions. Intelligent Systems Technologies and Applications 2016. AISC, vol. 530, pp. 3–19. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47952-1_1
6. Heisele, B.: Visual object recognition with supervised learning. IEEE Intell. Syst. **18**(3), 38–42 (2003)
7. Erhan, D., Szegedy, C., Toshev A., Anguelov D.: Scalable object detection using deep neural networks. In: CVPR (2014)
8. Bappy, J.H., Roy-Chowdhury, A.: CNN based region proposals for efficient object detection. In: IEEE International Conference on Image Processing (ICIP) (2016)
9. Yan, J., Yu, Y., Zhu, X., Lei, Z., Li, S.Z.: Object detection by labeling superpixels. In: CVPR (2015)

10. Blasiak, A.: A Comparison of image segmentation methods. Thesis in Computer Science (2007)
11. Pantofaru, C., Hebert, M.: A comparison of image segmentation algorithms. Carnegie Mellon University (2005)
12. Robinson, D.J., Redding, N.J., Crisp, D.J.: Implementation of a fast algorithm for segmenting SAR imagery. Scientific and Technical report, Defense Science and Technology Organization, Australia (2002)
13. Rosenberger, C., Chabrier, S., Laurent, H., Emile, B.: Unsupervised and supervised image segmentation evaluation. Adv. Image Video Segm. **29**(1), 365–393 (2006)
14. Zhang, Y.J.: A survey on evaluation methods for image segmentation. Pattern Recognit. **29**, 1335–1346 (1996)
15. Li, S., Wu, D.O.: Modularity-based image segmentation. IEEE Trans. Circ. Syst. Video Technol. **25**(4), 570–581 (2015)
16. Felzenszwalb, P., Huttenlocher, D.: Efficient graph-based image segmentation. Int. J. Comput. Vis. **59**(2), 167–181 (2004)
17. Rao, S.R., Mobahi, H., Yang, A.Y., Sastry, S.S., Ma, Y.: Natural image segmentation with adaptive texture and boundary encoding. In: Zha, H., Taniguchi, R., Maybank, S. (eds.) ACCV 2009. LNCS, vol. 5994, pp. 135–146. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12307-8_13
18. Browet, A., Absil, P.-A., Van Dooren, P.: Community detection for hierarchical image segmentation. In: Aggarwal, J.K., Barneva, R.P., Brimkov, V.E., Koroutchev, K.N., Korutcheva, E.R. (eds.) IWCIA 2011. LNCS, vol. 6636, pp. 358–371. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21073-0_32
19. Comaniciu, D., Meer, P.: Mean shift: a robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Mach. Intell. **32**(7), 1271–1283 (2010)
20. Soille, P.: Morphological Image Analysis: Principles and Applications. Springer, Berlin (2004). https://doi.org/10.1007/978-3-662-05088-0
21. Cour, T., Benezit, F., Shi, J.: Spectral segmentation with multiscale graph decomposition. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR, vol. 2, pp. 1124–1131. IEEE (2005)